

Tangible interfaces: when physical-virtual coupling may be detrimental to learning

Sébastien Cuendet
CRAFT - EPFL
Rolex Learning Center - Station 20
CH-1015 Lausanne - Switzerland
sebastien.cuendet@epfl.ch

Patrick Jermann
CRAFT - EPFL
Rolex Learning Center - Station 20
CH-1015 Lausanne - Switzerland
patrick.jermann@epfl.ch

Pierre Dillenbourg
CRAFT - EPFL
Rolex Learning Center - Station 20
CH-1015 Lausanne - Switzerland
pierre.dillenbourg@epfl.ch

Tangible user interfaces (TUIs) have been the focus of much attention recently in the HCI and learning communities. Although TUIs seem to intuitively offer potential to increase the learning experience, there have been questions about whether they actually impact learning positively. TUIs offer new ways of interactions and it is essential to understand how the design choices made for these new interactions affect learning. One element that is key in the learning process is how and when feedback is provided. In this article, we focus on the effect of co-located immediate process-level feedback on learning. We report the results of a study in which 56 participants used a TUI to complete tasks related to the training of spatial skills. Half of the students accomplished the tasks with immediate and co-located feedback from the system, while the other half of the students did not receive any feedback. Results show that participants who did not receive feedback manipulated less, reflected more, and in the end learned more than those who received feedback.

TUI, tangible, learning, feedback, coupling, dyna-linking

1. INTRODUCTION

Tangible User Interfaces are (TUIs) intuitively thought of as favoring learning, because they involve physical behaviors that accompany cognition, such as gesturing, physical movement and embodiment (Goldin-Meadow (2003); O'Malley & Stanton Fraser (2004)). TUIs also provide external representations of a problem or an object, which play a key role in problem solving and learning (Ainsworth (1999); Larkin & Simon (1987); Zhang & Norman (1994)) by helping the learner to make inferences or freeing up cognitive load to allow the learner to focus on the core of his task. Numerous examples of tangible interfaces for learning can be found in Shaer & Hornecker (2009). Marshall and colleagues have categorized the possible benefits brought to learning by tangibles in 4 categories: collaboration, accessibility, novelty of links, and playful learning (Marshall et al. (2007)).

However, many questions have been raised about the actual impact of TUIs on learning and there is still a lack of formal research on their benefits in terms of learning outcomes. Indeed, while many studies reported a higher engagement when using TUIs (e.g. Price & Rogers (2004)), the few controlled experiments comparing graphical user interfaces

(GUIs) and TUIs have shown no difference between GUIs and TUIs in learning performance, according to Marshall (2007). In the same article, the author questioned the real benefits of tangibles for learning, and suggested that more empirical investigation was needed to measure them.

Like others have argued in the past with other technologies, we argue that TUIs are neither activators nor inhibitors of learning per se. Instead, it is how they are used – and especially the cognitive activity related to the tangible interface – that are decisive on whether or not TUIs have a positive impact on learning. More precisely, we examine the information that is provided by the system in reaction to the learners' input, i.e. the *feedback* given by the system. Coupling the tangible input with a visual output is the cornerstone of the high usability and engagement of tangible interfaces. However, this continuous and immediate feedback might also act as a kind of prosthesis that substitutes to activating the targeted cognitive processes required for learning.

Small design variations, such as providing real-time feedback or not, can foster or prevent cognitive mechanisms that are beneficial for learning. Delayed feedback, provided when learners have completed

a task, gives the learners opportunities to fail and to trigger cognitive processes to repair or overcome their errors. While coupling increases usability, it may be detrimental to learning. To empirically assess this, we exposed carpenter apprentices to a tangible environment; some of them were provided with coupling while others were not. They completed tasks related to spatial skills training and learning was measured as the improvement of spatial reasoning skills specific to orthographic projections. Based on this set-up, we sought to examine the impact on the behavior and learning of participants generated by the presence or absence of coupling.

2. RELATED WORK

2.1. Multiple external representations (MER)

As mentioned in the introduction, one potential benefit of TUIs is that tangibles can be used as external representations, which have been shown to be powerful for learning. Larkin & Simon (1987) compared computational efficiency of sentential and diagrammatic representations and showed that the diagrammatic representation can allow the user to find information more efficiently. One advantage of MERs is that they allow the user to see how his action on one representation impacts a second representation, as explained by Ainsworth (2006). This is commonly referred to as dyna-linking and can be qualified as one kind of immediate feedback. Dyna-linking is assumed to reduce cognitive load, allowing the students to concentrate on their actions on the representations and their consequences. Price et al. (2009) showed that it can also encourage learners to try to understand concepts that are beyond their level of understanding by encouraging them explore more.

However, empirical data supporting the benefits of dyna-linking for learning are hard to find. One explanation for this is that dyna-linking allows students to explore the links between representations without reflecting or planning their problem-solving moves. In cases where the solution is shown in one of the representations, the learners could solve the problem by perceptually aligning the representations by trial-and-error. One way to counter this effect is to vary the feedback provided to learners depending on their performance: the better a learner gets, the less support he gets. This has been suggested before, for example by Price et al. (2009) who argued for the need to "specifically design learning activities that slow down interaction and promote opportunities for reflection to occur during calm periods at various points in the learning task". Do-Lenh et al. (2010) also observed that the users of their tangible system tended to

manipulate too much and not reflect enough. In a subsequent study (Do-Lenh (2012)) they introduced reflection tools in their tangible learning environment and showed that it had a positive impact on the learning of students.

2.2. Feedback

The debate about providing immediate versus delayed feedback is as old as education. However, novel interfaces revived this debate from a new perspective, as a tight coupling of input and output (i.e. immediate or continuous feedback) contribute to their success in terms of usability and engagement.

Hattie (1999) found that feedback is in the top 5 to 10 highest influences on achievement in learning. Hattie & Timperley (2007) further discovered through a meta-analysis that feedback is most powerful in the context of faulty interpretations, and not lack of information; when goals are specific and challenging but task complexity is low; when it provides information on correct rather than incorrect responses and when it builds on changes from previous trials. The timing of feedback is an important factor with regards to efficiency. In a meta-analysis of 53 studies, Kulik & Kulik (1988) showed that immediate feedback is beneficial at the task level but that at the process level, some delay is beneficial. Similarly, in a study involving high school students who completed a computer-based lesson, Clariana et al. (2000) found that delayed feedback was effective when greater degrees of processing were needed. Their results indeed showed that delayed feedback was most effective for difficult problems. The authors suggested that the reason for this was that more processing was needed for the difficult items and that delayed feedback encouraged more processing about the task.

Based on their findings, Hattie & Timperley (2007) developed a model for feedback where they distinguished four levels at which feedback can happen. It can be about a specific task. It can be directed at the process used to complete the task. It can focus on self-regulation. Finally, it can also be personal and directed straight at the person. In agreement with Kluger & DeNisi (1998), who found that feedback is most efficient when it is close to the task as opposed to close to the person, Hattie & Timperley (2007) found that this last type of feedback is too often unrelated to the task and does not impact learning positively. They also indicate that feedback on process is more effective for deeper understanding than feedback on the task. The distinction between feedback levels is also addressed by a control theoretical view of behavior. Accordingly, behavior is controlled by a hierarchy of

feedback loops at different levels of control (Lord & Levy (1994)).

The role of feedback and physical manipulation has been studied by Brasell (1987). Students used a microcomputer-based laboratory where a motion graph is updated in real-time as a consequence of physically moving a toy car. The author found that students understand graphs of motion better when they can physically experiment by controlling the object and when they get immediate rather than delayed feedback. In a similar vein, according to Holton (2010), one of the ways to facilitate conceptual understanding following the tenets of embodied cognition, consists of making the phenomenon “visible and manipulable”.

2.2.1. Terminology

Different authors use different terminologies to describe the various levels of behavior regulation. Following the terms of Anderson (Anderson (2002)) the real-time coupling of physical manipulation and system augmentations corresponds to the level of operations. Users are informed every second (or even more often) about the effect of their actions. The provision of feedback about whether a question was solved correctly corresponds to the level of unit-task. Feedback about performance of a series of questions corresponds to the level of task. In the terms of Hattie & Timperley (2007), coupling the users’ manipulations with augmentations in tangible environments is a form of process feedback and providing feedback about the correctness of the solutions is a form of task feedback. In this work, we will refer to the process and task level.

3. QUESTION

One of the features of tabletop TUIs is that they allow users to have external representations, and to have them co-located with their display (Price (2008)). Co-located displays increase the potential for dynalinking by making it possible to show to the user a direct *feedback* of his actions on top of the TUIs. Our specific question in this contribution is whether the provision of feedback at the process level is beneficial for learning. This is a first step towards understanding the level of behavior at which it is most appropriate to provide feedback and how this feedback should be provided.

4. METHOD

4.1. Technical setup

For this experiment, we used the Tinkerlamp, shown in Figure 1. The TinkerLamp is a tabletop environment developed at CRAFT (Zufferey et al.



Figure 1: The Tinkerlamp.

(2008)). It is composed of a camera and a projector directed at a tabletop via a mirror. The projection area, i.e. the playground for applications, is of dimension 70 by 55 centimeters. The lamp is able to detect tagged objects placed under it thanks to a tag tracking library similar to the ARTag library introduced by Fiala (2005), and can provide visual feedback through the projector.

4.2. Participants

Fifty-six male carpentry apprentices in their first year of their training participated in the experiment. The age of the participants ranged between 16 and 21 years, with a couple of participants in their late twenties. They had been exposed to and had performed some exercises with orthographic projections prior to the experiment. Apprentices completed the activities in pairs. Each pair was randomly assigned to one condition. In total, 14 pairs completed the experiment in the coupling condition and 15 in the no-coupling condition.

4.3. Environment and method

The experiment was conducted during a drawing class, in a classroom in which two Tinkerlamps were set up in the back. The apprentices came in pairs to one of the Tinkerlamps to participate in the experiment. The experiment was conducted over 12 days with four classes.

The apprentices passed tests before and after completing the activity (pre and post-tests). Both tests had 6 questions. There were three types of questions: (1) given the top view of an object, students had to choose the corresponding face and side views from 8 candidates (3 questions); (2) given the side view and face view of an object, students had to pick the corresponding top view from 6 candidates (2 questions); and (3), given the side view of an object, students had to draw the

corresponding face view. There was no limit of time for either the tests or the treatment.

4.4. Task

The tabletop display that students used is shown in Figure 2. It consists in two parts. The “Block zone” (gray square on the bottom right) is where the tangible objects are manipulated. The “Projection zone” (on the top left) contains two orthographic projections of the block: its face and side views. The “Projection zone” as well as the “Block zone” has a square grid with spacing of 2 centimeters for the purpose of orientation. A score is displayed on the top right. Instructions and feedback appear as pop-up windows.

The activities contained a total of 24 questions, for which the students did not have any time limit. Each question consisted of four steps. First, students had to place a block in the “Block zone” at a position and orientation that the system indicated by drawing its top view. Second, once the block was placed correctly, a movement was applied virtually to the object and the new position of the object was shown on the face and side views. Third, the students had to move the block to this new position indicated in the “Projection zone”. This transition from one position to the other had to be done with as few movements as possible. Fourth, when they were satisfied with their solution, students could check their solution by placing a specific token on the tabletop. If the solution was wrong, the correct solution was displayed as a projection of the top view in the “Block zone” for 20 seconds, and students were asked to place the block as indicated. After this time, or immediately if the solution was correct, the next question started from the last position.

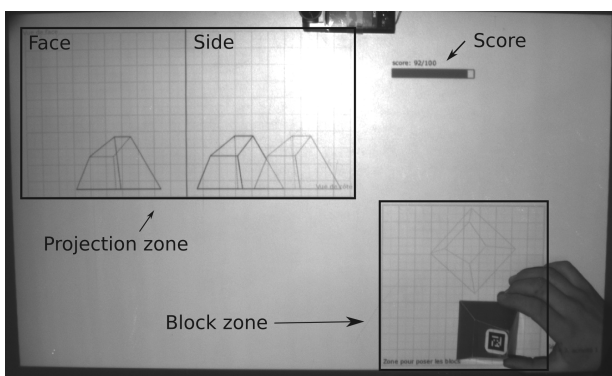


Figure 2: An snapshot of the tabletop during an activity. The object is placed in a square on the bottom right of the playground. Its face and side views are shown on the top left.

The three blocks used are shown in Figure 3; each of the blocks was used for 8 consecutive questions.

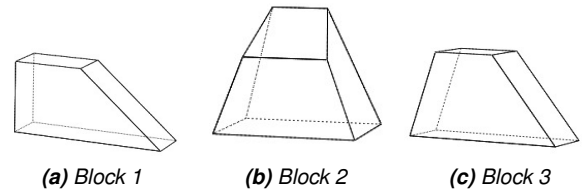


Figure 3: The shape of the three blocks. Block 1 and 3 were made out of wood and block 2 was made out of folded cardboard.

Besides finding the correct position, the students were also asked to maximize a score that was displayed permanently. The score started at a 100 and was inversely proportional to the length of the path done by the block. Moving the block between the starting position and the correct solution could be done without losing any points, but not taking the shortest path led to a loss of points. The score was meant to encourage the use of mental rotation and movement planning over trial-and-error. Finally, on half of the questions an animation of the transition from the starting to the new position of the block was shown before students starting moving the object.

4.4.1. Question types

For each question, the block could be moved either by a translation (on one or two axis), or by a translation combined with a rotation. Therefore, the movement applied to the block in one question could have from one two three dimensions. In the the rest of this work, we will refer to questions that had a movement with a translation only on the y-axis as “y” questions, to questions with a translation on x and y as “xy” questions, etc. Note that in a “y” question the side view changes and the face view stays the same, that it is the opposite for an “x” question, and that for “xy” questions, both the face and the side view change.

4.5. Conditions

There were two conditions: *coupling* and *no-coupling*. The only difference between the two conditions was that in the coupling condition, the tangible representation of the block was dynamically linked to the virtual representations shown on the face and side views. In other words, students in the coupling condition could see in real-time on the two orthographic projections the effect of moving the tangible block.

In terms of feedback, the students in the coupling condition were provided with immediate feedback at the process level whereas students in the no-coupling condition did not receive any process feedback. Both conditions offered immediate and delayed feedback at the task level. The immediate

feedback was given to students as "correct", "almost correct", and "wrong" when they finished a question. The delayed feedback appeared at the end of each series of 8 questions and recapitulated the result of each question (right/wrong) and, for correct questions, the score.

4.6. Statistical testing

The statistical tests were made using ANOVAs on linear models. Repetitions were taken into account using mixed effect models when needed.

5. RESULTS

5.1. Learning gain

There was an overall learning gain independently of the experimental condition (paired-sample t-test, $t(55)=2.28$, $p<0.05$). A test on each separate condition further shows that students in the no-coupling condition improved between the pre and the post-test ($t(28)=2.21$, $p<0.05$), whereas the improvement in the coupling condition is not statistically significant ($t(26)=1.07$, $p>0.05$). Figure 4 shows the means of the pre and post-test results for each of the condition together with the confidence interval. Students in the coupling condition performed slightly higher in the pre-test, whereas in the post-test the scores are almost identical.

5.1.1. Answer types

To better understand what caused the overall improvement, we distinguished between eight types of answers. The binary coding scheme for the answer type is as follows: <orientation is correct><face view is correct><side view is correct>. So for example, the 110 signature means that the orientation and the face view were correct, but the side view was not. Figure 5 shows the ratio of answers per answer type between the pre and post-test for each condition. The ratio takes into account the (minor) variations in the distribution of answer types between the pre and post-test. There is no ratio for the signature 101 because no answer in the pre-test had this signature.

In both conditions, the number of correct answers (111) increased, with a slightly larger increase for the no-coupling condition. All answer types for which the face view was wrong (*0*) were equally or less frequent in both conditions, except for the 100 type that increased slightly in the coupling condition. Furthermore, the number of 110 answers were the ones that increased the most.

However, there were some differences between the two conditions. For example, while the 110 answers increased the most in both conditions,

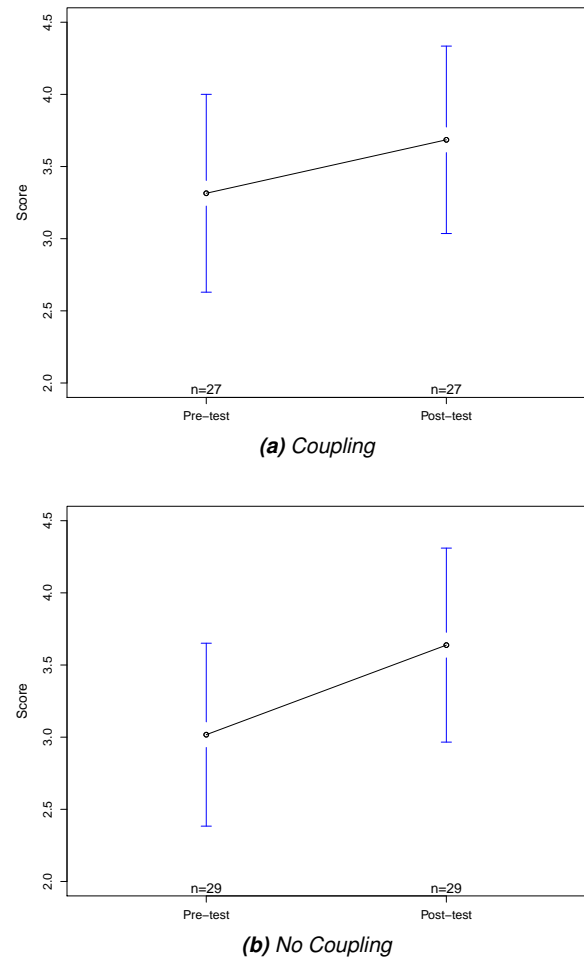


Figure 4: Means of the pre- and post-test scores for the two conditions.

the increase was twice as strong in the coupling case. In the coupling condition, the number of times that students selected the answers for which the rotation was correct (1**) increased in the post-test. Although we cannot compute a relative change for the improvement of the answers with a correct side view (101), we notice that for the coupling condition this answer was provided 8 times by 27 students (0.29) compared to 11 times by 29 students (0.38) for the no coupling condition. This might suggest that students in the coupling condition improved their ability to identify a correct orientation and positioning in the face view, but not in the side view.

5.2. Score evolution during treatment

The percentage of correct answers increased during the treatment in the no-coupling condition ($p<0.05$, $F[1,358]=8.20$) but not in the coupling condition, as depicted in Figure 6, which shows the percentage of questions solved correctly during the treatment for each of the conditions. This must however be taken with some care since the apparent ceiling effect in the coupling condition might explain part of it.

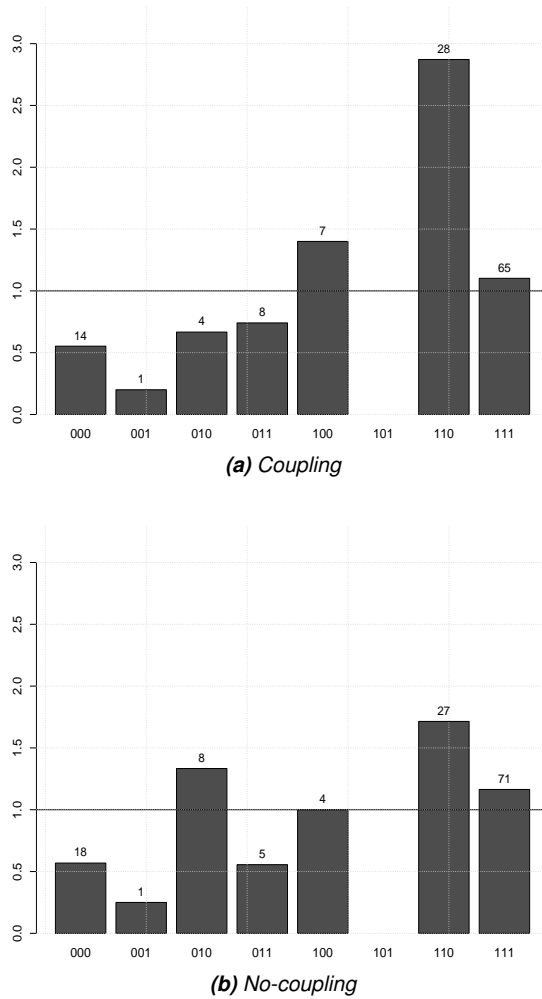


Figure 5: Change of proportion in the type of answer between pre-test and post-test. The ratio takes into account the distribution of the answer types. The horizontal line indicates the limit between an increase and a decrease. The numbers on the bar indicate the number of occurrences of this answer type in the post-test.

For score evolution, a similar pattern was observed only for the question type that involved two translations without rotation (Figure 7). In this case, there is an interaction effect between the condition and the evolution in the time ($F[1,47]=7.2, p<0.05$).

5.3. Link between the treatment score and the post-test score

The pre and post-test were taken individually, whereas the treatment was performed in pairs. When comparing treatment with the test scores, the test score of a group is computed as the mean of the test score of each of its two members. The performance of groups in the no-coupling condition during the treatment is related to their score at the post test ($r=0.65, t[13]=3.15, p<.001$). There is no such relationship for the groups in the coupling condition ($r=0.35, t[12]=1.28, p>.05$). The fact that

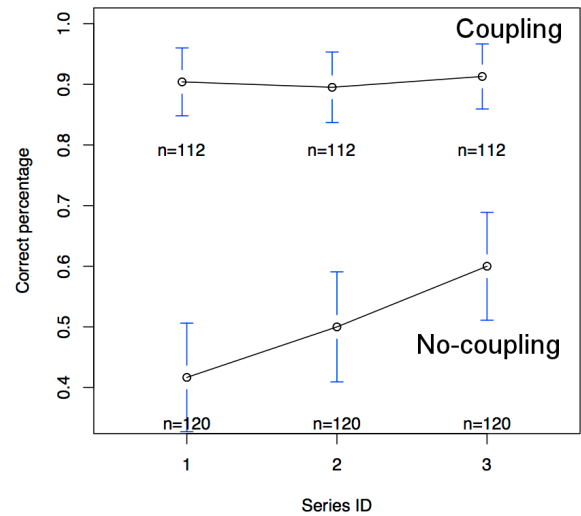


Figure 6: Results by series, which each contained eight questions. All questions in one series used the same block, but three different blocks were used for each of the series.

there is no relationship between the treatment score and the pre-test ($r=0.04, t[27] = 0.23, p>.05$) shows that the higher improvement during the treatment of no-coupling students and their higher performance in the post-test is not due to a prior bias. In other words, providing immediate feedback removed the correlation between the performances during the treatment and the post-test.

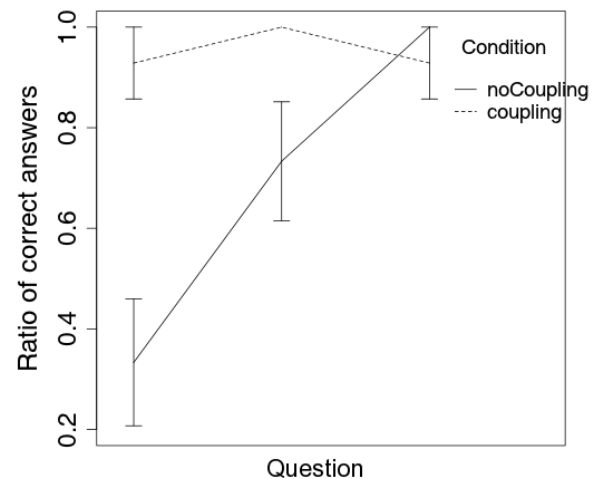


Figure 7: The interaction effect between condition over time for the "xy" type of questions, i.e. the questions that had one translation on each axis (x and y axis), but no rotation.

5.4. Behavioral differences during treatment

5.4.1. Waiting time and speed of move

Students in the coupling condition completed a question in 67 seconds in average, to compare with 88 seconds in the no-coupling condition. In the coupling condition, students waited in average 8 seconds before moving the block for the first time versus 14 seconds in the no-coupling condition. In the coupling condition, from the moment a question started and the moment they checked their solution, students moved the block for about 10% of the time, twice the percentage of the no-coupling condition. Similarly the average moving speed was larger in the no-coupling condition.

5.4.2. Path analysis

We categorized the movements of the blocks along a path into three types. A movement was *correct* if the block was moved closer to the correct solution on both x and y axes. A movement was *half-wrong* if the block ended up being closer to the solution only on one dimension. Finally, a movement was *wrong* when the block was moved in the wrong direction along both dimensions. The relative amount of each of these categories of movements is shown in Figure 8. The ratio of correct movements was significantly higher for the coupling condition. In the no-coupling condition, this ratio was approximately the same as the ratio of half-wrong movements.

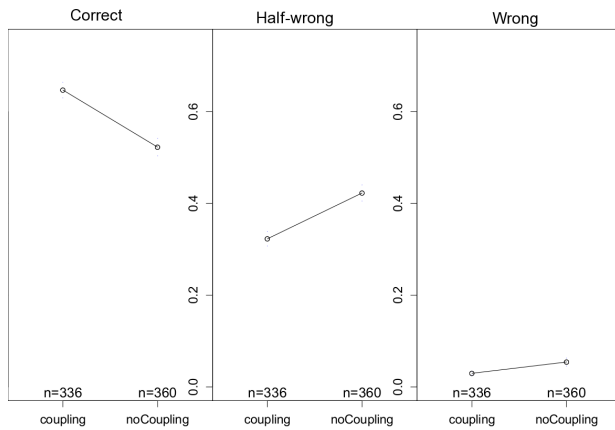


Figure 8: The ratio of correct, wrong, and half-wrong movements by condition.

Furthermore, the diversity of the paths among the groups was different across the two conditions. As can be observed in Figure 9, the paths chosen by students in the coupling condition were similar to each other. However, in the no-coupling condition, the groups chose very different paths and often ended at an incorrect position. One can also observe the strategies adopted: in the no-coupling condition, students who found the correct position did so by decomposing the movement first on the y-axis, and

then on the x-axis. In the coupling condition, there were three successful strategies: decomposing on both axes, or going straight in a diagonal line. The latter is the optimal strategy, but also the most difficult one because it implies moving along the two axes at the same time and therefore controlling two factors simultaneously.

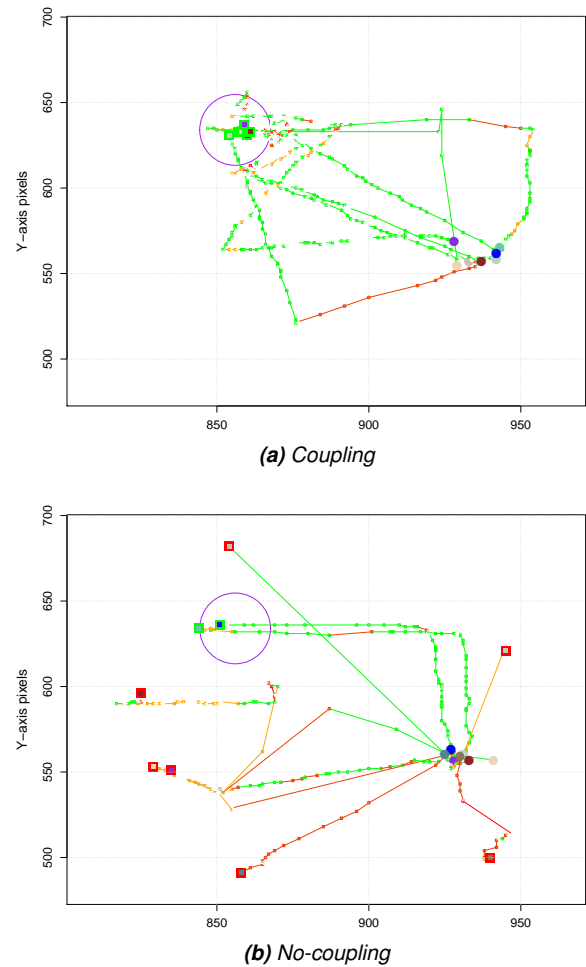


Figure 9: Detail of the path for one question. The purple circle indicates the correct solution. The full circles indicate the starting positions, and the squares the arrival points of the task. The color of the square indicates whether their solution was correct (green) or wrong (red). Each point represents a time lapse of 1 second. The red segments indicate a move for which both the x and y move were wrong; segments are orange when one of the two directions was wrong; the green segments correspond to correct moves.

Similarly to the number of correct answers to a question, the difference in the path pattern was more distinct for some types of questions. There was a significant interaction effect between the condition and the type of questions for the "y" question type. This interaction effect came from the fact that no-coupling students significantly increased their ratio of correct movement for this type of questions, while

the ratio of correct movement decreased for coupling students.

5.4.3. Variations among question types

It appeared that some types of questions were more difficult to solve than others. We grouped the questions by the number of translations (first number) and whether or not a rotation was involved (0 means no rotation, 1 means there was a rotation). Figure 10 shows the average ratio of correct answers during the treatment for four categories of questions for both conditions. There is an interaction effect between the condition and the question type, indicating that some questions were harder to solve in one condition. In the coupling condition, all questions had a similarly high score. On the contrary, in the no-coupling condition, there were significant differences between the different types of questions: students scored significantly lower on questions involving a rotation than on those without one ($F[1,344]=33.6, p<0.05$). The number of translations did not affect the result.

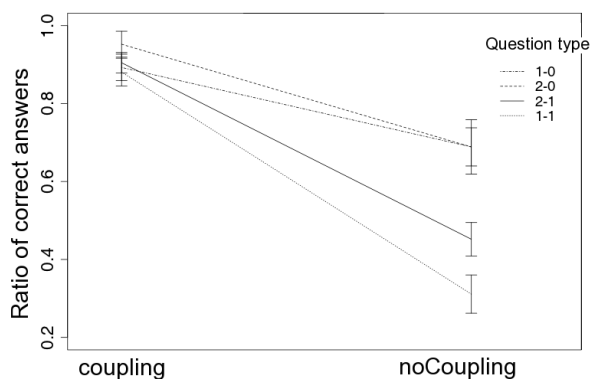


Figure 10: Ratio of correct answers given during the treatment, by type of questions and by condition. The coding scheme for the type of questions is <translation>-<rotation>. For translations, the number indicates along how many axis the translation occurred; for the rotation, 1 means there was a rotation, 0 means there was not. So for example, 2 – 0 means that there were translations on both x and y but that there was no rotation.

5.4.4. Effect of showing the transition

Showing the transition did not have any effect on the percentage of correct answers given overall. It seemed to increase the performance when solving questions with a rotation in the no-coupling condition, but the effect proved non-significant.

6. DISCUSSION

The type of feedback provided in the coupling condition was at the process level. It nevertheless had an impact on students' behavior and learning both at the process and the task level.

6.1. Differences at the process level

At the process level, the analysis of the path patterns revealed behavioral differences between the two conditions. Students in the coupling condition manipulated more and did not wait as long as no-coupling students before performing actions. Additionally, once they started moving the block, they moved it slower than no-coupling students. Since there is no reason to think that students in the coupling condition were naturally more prone to action, this suggests that the dyna-link between the tangible blocks and the virtual representations encouraged students to rapidly dive into action without much beforehand thinking, as if proceeding in a trial-and-error fashion. Students in the coupling condition should not be blamed for their lack of reflection: they were told to move the block to a new position using the shortest possible path. They quickly discovered that by slightly moving the block around the starting position they could use the feedback to infer the correct direction.

However, due to the lack of immediate feedback, students in the no-coupling condition reflected more before acting. They also ended up with many different solutions, as indicated by the heterogeneity of the paths. Note that this could be used in the future to generate a confrontation between the groups. This could improve learning, since confronting diverging opinions is known to favor learning Doise et al. (1991). The path patterns in the no-coupling condition showed a higher rate of half-wrong segments, hinting that the students in this condition decomposed the movement into sequential movements along the two axes.

Finally, the effect of the score was higher than we expected. Students were serious about trying not to waste points. This was reflected by very small movements and a trial-and-error approach in the coupling condition and by a longer thinking time in the no-coupling condition. Including the score as a performance metric was a deliberate design choice for this experiment, with the goal of restraining the coupling students. The same experiment without the score would probably have led to different results.

6.2. Differences at the task level

The differences at the process level between the two conditions were reflected at the task level in the learning outcome. The no-coupling condition, in which students reflected more and manipulated less, led to a learning gain. Moreover, the treatment performance was positively correlated with the post-test score for the no-coupling students, but not for the coupling students. This suggests that the higher level of manipulation involved in the coupling condition

was not beneficial for learning. This is in agreement with Price et al. (2009) and further questions some of the general benefits of manipulation claimed by the embodied cognition approach (Holton (2010); Schwartz & Holton (2000)). An explanation for this is that the task at hand in this study was not only sensori-motor but also included the ability to elaborate a mental image of the 3D movements.

The type of answers given by students indicated that it was easier for students to find the right position on the x-axis than on the y-axis. This is in agreement with the literature on spatial skills that showed that the greater the mental rotation the harder the representation (e.g. Flusberg & Boroditsky (2011)). The difference in the answers given between the pre-test and the post-test also revealed that the no-coupling students improved more evenly than coupling students, whose improved at identifying the correct orientation and the correct position on the x-axis, but not the correct position on the y-axis.

In all the groups, both members showed a high engagement and actively participated in the task. In rare occasions, some groups used the tangible setting to bypass the need to do mental rotation, by positioning one member around the table so that he could see the side view as a face view. This strategy defines the roles of viewer and manipulator, and its impact on learning for each of the two members remains to be studied.

Questions involving a rotation were harder to solve in the no-coupling condition, whereas all questions were of the same difficulty in the coupling condition. However, showing the transition seemed to help the no-coupling students when solving rotation-related questions. Although this difference was not significant, this indicates that animations could be helpful to understand the impact of a rotation on orthographic projections. At least some students noticed and used the display of the transition, saying things such as "You just have to follow the movement that it is showing".

7. CONCLUSIONS AND FUTURE DIRECTIONS

The purpose of this study was to observe in depth the impact on learning and students' behavior when changing the type of feedback given at the process level in a TUI environment. 56 students were split into two conditions, one with coupling between the tangible and the virtual representations, and the second without. Results showed that dyna-linking led to a higher degree of manipulation and a lower degree of reflection. It made the task easier and students in this condition solved more questions correctly during the treatment. However

those students did not improve during the treatment, as opposed to students who were not provided with the coupling. Moreover, a pre-test/post-test assessment showed that students with the dyna-linking did not improve significantly, while the others did.

At a time where the benefits of TUIs for learning are questioned, these results support the arguments made by Marshall et al. (2007) that tangibles can improve learning, but that special attention needs to be given to the design of the tangible environment. The presence or absence of physical-virtual coupling can lead to important differences in students' behavior and subsequently, on learning outcome. Because of its increased usability physical-virtual coupling is often the default in TUI environments, possibly at the expense of learning. This is a small piece of the puzzle and much has still to be learned as to how TUIs can be used efficiently for learning.

An interesting future direction would be to explore how to use different types of feedback in the same activity, but with a different timing. Since our results show that immediate feedback makes the task easier, it could be useful to provide it when exploring a new subject to lower the entry-level threshold. Since it does not promote reflection, this type of immediate feedback could be gradually diminished as students progress into the activity. Exploring the influence of other kinds of feedback on TUIs, such as feedback at the self-regulating level is another direction.

8. ACKNOWLEDGEMENTS

Thanks to Engin Bumbacher for his careful proof-reading and his many helpful suggestions, and to MM. Ogay, Wuillamoz, and Rüfenacht for opening their classrooms to us. This work is supported by the Federal Office for Professional Education and Technology (OPET).

REFERENCES

- Ainsworth, S. (1999), 'The functions of multiple representations', *Computers & Education*.
- Ainsworth, S. (2006), 'DeFT: a conceptual framework for considering learning with multiple representations', *Learning and Instruction* **16**(3), 183–198.
- Anderson, J. (2002), 'Spanning seven orders of magnitude: A challenge for cognitive modeling', *Cognitive Science* **26**(1), 85112.
- Brasell, H. (1987), 'The effect of real-time laboratory graphing on learning graphic representations of

- distance and velocity', *Journal of Research in Science Teaching* **24**(4), 385–395.
- Clariana, R., Wagner, D. & Roher Murphy, L. (2000), 'Applying a connectionist description of feedback timing', *Educational Technology Research and Development* **48**(3), 5–22.
- Do-Lenh, H. S. (2012), Supporting Reflection and Classroom Orchestration with Tangible Tabletops, PhD thesis, EPFL, Lausanne.
- Do-Lenh, S., Jermann, P., Cuendet, S., Zufferey, G. & Dillenbourg, P. (2010), Task performance vs. learning outcomes: a study of a tangible user interface in the classroom, in 'Proceedings of the 5th European conference on Technology enhanced learning conference on Sustaining : from innovation to learning and practice', EC-'10, Springer-Verlag, Berlin, Heidelberg, p. 7892.
- Doise, W., Deschamps, J. & Mugny, G. (1991), *Le conflit socio-cognitif*, Armand-Colin, Paris.
- Fiala, M. (2005), Artag, a fiducial marker system using digital techniques, in 'Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on', Vol. 2, pp. 590–596 vol. 2.
- Flusberg, S. & Boroditsky, L. (2011), 'Are things that are hard to physically move also hard to imagine moving?', *Psychonomic bulletin & review* **18**(1), 158164.
- Goldin-Meadow, S. (2003), *Hearing gesture: How our hands help us think.*, Harvard University Press.
- Hattie, J. (1999), 'Influences on student learning', *Inaugural lecture given on August 2*, 1999.
- Hattie, J. & Timperley, H. (2007), 'The power of feedback', *Review of Educational Research* **77**(1), 81–112.
- Holton, D. (2010), Practical implications of embodied cognition theory and enactivism for instructional design, Denver, Colorado, United States.
- Kluger, A. N. & DeNisi, A. (1998), 'Feedback interventions: Toward the understanding of a Double-Edged sword', *Current Directions in Psychological Science* **7**(3), 67–72.
- Kulik, J. A. & Kulik, C. C. (1988), 'Timing of feedback and verbal learning', *Review of Educational Research* **58**(1), 79–97.
- Larkin, J. H. & Simon, H. A. (1987), 'Why a diagram is (sometimes) worth ten thousand words', *Cognitive Science*.
- Lord, R. G. & Levy, P. E. (1994), 'Moving from cognition to action: A control theory perspective', *Applied Psychology* **43**(3), 335–367.
- Marshall, P. (2007), Do tangible interfaces enhance learning?, in 'Proceedings of the 1st international conference on Tangible and embedded interaction', ACM, pp. 163–170.
- Marshall, P., Rogers, Y. & Hornecker, E. (2007), Are tangible interfaces really any better than other kinds of interfaces?, in 'CHI'07 workshop on Tangible User Interfaces in Context & Theory'.
- O'Malley, C. & Stanton Fraser, D. (2004), Literature review in learning with tangible technologies, Technical report, FutureLab.
- Price, S. (2008), 'Towards a framework for investigating tangible environments for learning', *International Journal of Arts and Technology* **1**(Special Issue on Tangible and Embedded Computing), 351–368.
- Price, S. & Rogers, Y. (2004), 'Let's get physical: the learning benefits of interacting in digitally augmented physical spaces', *Comput. Educ.*
- Price, S., Falco, T. P., Sheridan, J. G. & Roussos, G. (2009), The effect of representation location on interaction in a tangible learning environment, in 'Proceedings of the 3rd International Conference on Tangible and Embedded Interaction', ACM Press, p. 85.
- Schwartz, D. L. & Holton, D. L. (2000), 'Tool use and the effect of action on the imagination', *Journal of Experimental Psychology. Learning, Memory, and Cognition* **26**(6), 1655–1665.
- Shaer, O. & Hornecker, E. (2009), 'Tangible user interfaces: Past, present and future directions (preprint)', *Found. Trends Hum.-Comput. Interact.* **3**(1-2), 23–27.
- Zhang, J. & Norman, D. A. (1994), 'Representations in distributed cognitive tasks', *Cognitive Science* **18**(1), 87–122.
- Zufferey, G., Jermann, P. & Dillenbourg, P. (2008), A Tabletop Learning Environment for Logistics Assistants: Activating Teachers, in D. Cunliffe, ed., 'Proceedings of the Third IASTED International Conference Human-Computer Interaction', ACTA Press, pp. 37–42.